

EJERCICIO PRÁCTICO 1: MANEJANDO TABLAS DE DATOS EN R

CONTEXTO

En este primer ejercicio práctico comenzaremos a trabajar con el entorno de programación R y a experimentar las dificultades de trabajar con datos, mientras vamos reconociendo los conceptos estudiados.

OBJETIVOS DE APRENDIZAJE

1. Resolver dificultades para trabajar con datos entregados por terceras personas.
2. Cargar datos y usar las herramientas de R para hacer análisis básicos.

ÉXITO DE LA ACTIVIDAD

1. El equipo consigue el conjunto de datos solicitado en formato CSV para ser trabajado con R.
2. El equipo carga el conjunto de datos (importa) en el entorno R.
3. El equipo consigue trabajar con los datos cargados aplicando herramientas disponibles en R.

ACTIVIDADES

1. Descargar el archivo con datos sobre los nuevos casos diarios de pacientes sintomáticos diagnosticados con Covid-19 (EP01 Datos Covid.csv) desde UVirtual.
2. Importar el conjunto de datos al entorno de programación R y responder las siguientes preguntas:
 - ¿Qué variables se han cargado?
 - ¿Qué tipo tiene cada una de estas variables?
 - ¿Qué escala parecen tener estas variables?
3. Asegurar que las variables tienen el tipo correcto en la tabla de datos (data frame) cargada.
4. Utilizando herramientas de R (se recomienda usar el paquete dplyr) responda las preguntas asignadas a su grupo.
5. Subir la respuesta (script R) de lo que haya alcanzado a hacer al término de la clase, con nombre "EP01-respuesta-grupo-i", donde i es el número de grupo asignado.
6. Calcular $j = (i + 4) \bmod n + 1$, donde n es la cantidad de grupos; descargar la respuesta del grupo j y aplicar la rúbrica publicada para este ejercicio práctico.
7. Subir el resultado de la evaluación con el nombre "EP01-resultado-grupo-j".

PREGUNTAS

Grupo 1:

1. ¿Qué día se produjo el mayor número de casos con síntomas en la región del Maule entre el 01-oct-2020 y el 31-may-2021?
2. ¿Cuál fue el total de casos con síntomas para cada mes de este periodo?

Grupo 2:

1. ¿Qué día se produjo el mayor número de casos con síntomas en la región de Tarapacá entre el 01-mar-2021 y el 31-ago-2021?
2. ¿Cuál fue el total de casos con síntomas para cada mes de este periodo?

Grupo 3:

1. ¿Qué día se produjo el mayor número de casos con síntomas en la región Metropolitana entre el 01-ene-2021 y el 31-dic-2021?
2. ¿Cuál fue el total de casos con síntomas para cada mes de este periodo?

Grupo 4:

1. ¿Qué día se produjo el mayor número de casos con síntomas en la región de Coquimbo entre el 01-may-2020 y el 31-oct-2020?
2. ¿Cuál fue el total de casos con síntomas para cada mes de este periodo?

Grupo 5:

1. ¿Qué día se produjo el mayor número de casos con síntomas en la región del Maule entre el 01-mar-2021 y el 31-ago-2021?
2. ¿Cuál fue el total de casos con síntomas para cada mes de este periodo?

Grupo 6:

1. ¿Qué día se produjo el mayor número de casos con síntomas en la región de Ñuble entre el 01-nov-2020 y el 31-may-2021?
2. ¿Cuál fue el total de casos con síntomas para cada mes de este periodo?

Grupo 7:

1. ¿Qué día se produjo el mayor número de casos con síntomas en la región de Los Ríos entre el 01-jun-2020 y el 31-dic-2020?
2. ¿Cuál fue el total de casos con síntomas para cada mes de este periodo?

Grupo 8:

1. ¿Qué día se produjo el mayor número de casos con síntomas en la región de Los Lagos entre el 01-ene-2021 y el 30-jun-2021?
2. ¿Cuál fue el total de casos con síntomas para cada mes de este periodo?

Grupo 9:

1. ¿Qué día se produjo el mayor número de casos con síntomas en la región Metropolitana entre el 01-abril-2021 y el 30-ago-2021?
2. ¿Cuál fue el total de casos con síntomas para cada mes de este periodo?

Grupo 10:

1. ¿Qué día se produjo el mayor número de casos con síntomas en la región de Valparaíso entre el 01-oct-2021 y el 28-feb-2022?
2. ¿Cuál fue el total de casos con síntomas para cada mes de este periodo?

CRITERIOS DE EVALUACIÓN

- Obtienen un archivo con formato CSV preparado para ser cargado en un data frame de R de forma eficiente
- Mencionan explícitamente todas las variables presentes en el archivo CSV, reconociendo correctamente sus tipos y escalas

- Escriben código R que carga el archivo CSV y asegura que los tipos de variables se reflejen en el data frame resultante
- Obtienen correctamente la fecha en que ocurre el dato buscado analizando únicamente el periodo de fechas especificado y la región de Chile indicada
- Obtienen correctamente los valores solicitados agrupando mensualmente el periodo de fechas indicado y para la región de Chile indicada
- El script R es ordenado, está bien indentado, no contiene sentencias espurias y está bien comentado, con buena ortografía (<5 errores) y buena redacción, usando vocabulario propio de la disciplina y el contexto del problema
- Escribe con buena ortografía y redacción, usando vocabulario propio de la disciplina y el contexto del problema.